Journal of Solid State Chemistry **145,** 150–158 (1999)
Article ID jssc.1999.8233, available online at http://www.idealibrary.com on IDEAL®

Structure Maps of Close-Packed Layered Compounds and DNA

J. Hauck,* D. Henkel,* M. Klein,† and K. Mika*

*IFF and †IBT, Research Center Jülich, D-52425 Jülich, Germany

Received October 30, 1998; in revised form February 18, 1999; accepted February 19, 1999

Close-packed layered compounds such as $Fe_xGa_2S_{3+x}$ and $N_xC_3Al_{4+x}$ ($x = 0$–4) are analyzed as chains of covalent-bonded S=Ga(–S–Fe)$_x$–S–Ga=S or Al≡C(–Al=N)$_x$–Al=C=Al(–N=Al)$_{x'}$–C≡Al with delocalized bonds. The rules for valence compounds allow *RM, R(MR)$_n$* ($n = 1, 2, 3$), *RMMR,* and different combinations and substitutions. The conformation of the chains is characterized by the Jagodzinski–Wyckoff symbols *c* and *h* and a similar notation *s* and *t* at the occupation of tetrahedral sites. The interactions between the different atoms are analyzed by structure maps with the self-coordination numbers of nearest and next-nearest neighbors as parameters. A similar analysis of 18 different DNA or RNA reveals four different groups within different fields of the structure map similar to solid solutions or nonperiodic polytypes of layered compounds. © 1999 Academic Press

Key Words: **structure map; linear valence compounds; layered compounds; DNA.**

1. INTRODUCTION

All atoms of valence compounds either accept or provide valence electrons to obtain a stable octet configuration $n\ s^2np^6$; i.e., all *s* and *p* orbitals are completely filled or completely empty. The octets of the atoms can be completed in two different ways, as outlined for different covalent and ionic chains in Tables 1a. Electrons are shared between the atoms for covalent interactions or are donated by the metal atoms to the nonmetal atoms with increased electronegativity for ionic interactions (1, 2). According to Pauling the degree of ionicity of a bond is related to the electronegativity difference between the atoms (1–3). Structure maps (4, 5) are a different approach to analyzing interactions between atoms, as is outlined for linear chains of *a* and *b* atoms (or structural units). The different structures are characterized by the self-coordination numbers T_i of the minority component *a* with other *a* atoms (or units) in the chain with a_xb_y composition ($r = y/x \geq 1$). Each *a* or *b* in a chain has a maximum of $T_i^{\max} = 2$ nearest, next-nearest, and third-nearest neighbors ($i = 1$–3). In long chains the different T_i values of a species are averaged. The structures are characterized by $T_1\ T_2\ T_3$; *r* and are plotted in $T_1\ T_2$ maps for constant *r* or in an $\alpha_1\ \alpha_2$ map for different *r* values (Fig. 1). The short-range order parameters ($-1 \leq \alpha_i \leq 1$) can be obtained from T_i values by $\alpha_i\ T_i^{\max} = T_i - (T_i^{\max} - T_i)/r$. All structures at $r = 1$ composition are within a triangle with the three structures *ab*, a_2b_2, and $a_\infty b_\infty$ at corners and a random distribution in the center ($\alpha_i = 0$, $T_i = 2/(r+1)$). The borderline is varied in particular for low α_1 and α_2 values at increasing *r* with $-\alpha_i = 1/r$. The structures at the corners can be correlated with repulsive interactions (*ab*), attractive interactions (a_2b_2), or segregation ($a_\infty b_\infty$). The interactions are varied for structures at the borderlines, which can be considered combinations of structural units. The $ab_2a_2b_2$ structure at the upper right border, for example, is a combination of a_2b_2 and ab_2 units. The α_i values of *a* and *b* atoms are identical for binary compounds a_xb_y. The chains of the present investigation, containing three or four constituents such as the atoms *M*, *N*, and *R* in layered compounds or the four different bases in nucleic acids, exhibit up to three or four different points in the structure map if the self-coordination numbers and concentrations are different.

In close-packed layered compounds $M_xN_yR_z$ (Table 1) the *M*, *N*, and *R* atoms contribute different numbers of electrons to the bonding of the neighboring atoms. The formation of close-packed structures, many as polytypes with different stacking, indicates weak interactions. We suggest delocalized covalent bonds to the three atoms of neighboring layers similar to bonds in aromatic compounds or CO_3^{2-} groups and to dipole–dipole bonding between neighboring *R* atoms. The bond strength shown for the chain will be reduced by the delocalization. The present investigation shows a correlation between the sequences of layers and the stacking sequences and can explain some defects such as vacancy formation or a disorder of *M* and *N* atoms. The distance $\Delta z\ c/a$ between layers, which is ≈ 0.8 for close-packed metal atoms (N form), is reduced to 0.2–0.8 at the occupation of interstitial sites (I form) or to 0.2 for ordered body-centered alloys (B form).

The basic components of DNA are alternating units of 2′-deoxyribose (D) and phosphate groups (P) forming a DP chain with phosphodiester bonds between the 5′ position of

0022-4596/99 $30.00

TABLE 1a

Composition $M_xN_yR_z$ and Different Types of Chains *RM*, $(RM)_nR$ ($n=1-3$), and *RNRM* Which Can Be Formed by Atoms with One, Two, or Three Bonds, Number of Valence Electrons of *M*, *N*, and *R* Atoms, and Some Experimental or Theoretical Examples (in Brackets)

$M_xN_yR_z$	$e_Me_Ne_R$	Covalent chain	Ionic chain
RM	17	I–Cu	[I] Cu
RM	26	S=Zn or –S–Zn–	[S] Zn
RM	35	N≡Al or –N=Al– or =N–Al=	[N] Al
RM	44	–C≡Si– or =C=Si=	[C] Si
RMR	$0\,2\,7_2$	I–Cd–I	[I] Cd [I]
RMR	$0\,6\,1_2$	(Cu–S–Cu)	Cu [S] Cu
RMR	$0\,4\,6_2$	S=Sn=S	[S] Sn [S]
RMR	$0\,4\,2_2$	(Mg=C=Mg)	Mg [C] Mg
$(RM)_2R$	$0\,3_2\,6_3$	S=In–S–In=S	[S] In [S] In [S]
$(RM)_2R$	$0\,5_2\,2_3$	(Mg=N–Mg–N=Mg)	Mg [N] Mg [N] Mg
$(RM)_3R$	$0\,4_3\,5_4$	N≡Th–N=Th=N–Th≡N	[N] Th [N] Th [N] Th [N]
$(RM)_3R$	$0\,4_3\,3_4$	Al≡C–Al=C=Al–C≡Al	Al [C] Al [C] Al [C] Al
RNRM	$2\,4\,5_2$	N≡Th–N=Mg or –N=Th=N–Mg–	[N] Th [N] Mg
RNRM	$6\,4\,3_2$	Al≡C–Al=S or –Al=C=Al–S–	Al [C] Al [S]

TABLE 1b

Different Elements with e_i Valence Electrons in the Order of Increasing Size [10]

e_i	
1	Cu Ag Li
2	(Ni Co Fe Mn) Zn (Ge) Cd Hg (Sn Mg Pb Th)
3	(Ni Fe Cr) Al Ga In (Bi)
4	C Si (Pt Mo W Ta Nb) Sn (Zr)
5	N P As
6	S Se Te
7	Cl Br I

Note. The elements shown in brackets contain also other electrons. Zn, Al, Ga, and N atoms (underlined) occupy tetrahedral, Mn, In, C, and trivalent Fe (dashed underlined) tetrahedral or octahedral sites.

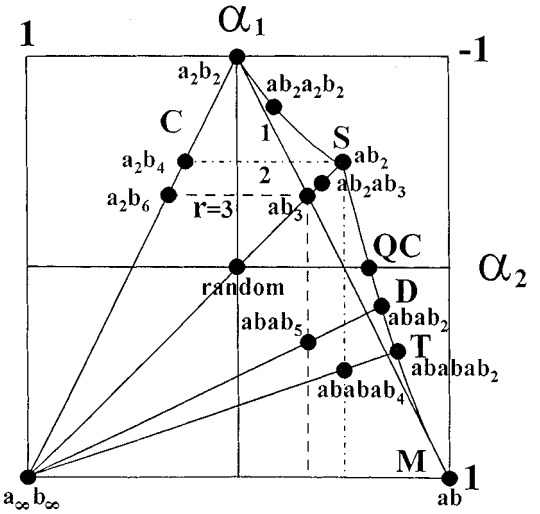

FIG. 1. Structure map of one-dimensional chain of a and b units with short-range order parameters $1 \geq \alpha_{1,2} \geq -1$, which are obtained from the self-coordination number T_1 of nearest and T_2 of next-nearest neighbors by $2\alpha_i = T_i - (2 - T_i)/r$, where $r \geq 1$ is the ratio of a and b units. The maximum range of the map varies with r. The different lines C, S, D, and T and the point M contain clusters (C) of a_2, a_3, etc., single units (S) ab_2, ab_3, etc., double units (D) $abab_2$, $abab_5$, triple units (T) $abababb_2$, $abababb_4$, or multiple units (M) ab of a. QC is obtained for a layered quasi-crystal.

one deoxyribose and the 3′ position of the next deoxyribose (6–9). One of the four bases adenine (a), thymine (t), cytosine (c), and guanine (g) is attached to the 1′ position of each deoxyribose. Each a is linked to a t of the neighboring chain by two hydrogen bonds, and each g is linked to a c by three hydrogen bonds. The a and g units are larger than the t and c units. The complementary strands are twisted around each other in the now familiar "double helix," with the bases in the center and the hydrophilic deoxyribose (or ribose in RNA) and phosphate residues on the outside. The exact DNA structure is influenced by several factors such as the sequence of bases and the interaction with the surrounding water. The Watson and Crick structure (B form) is a right-handed helix with 10 base pairs (bp) per turn and a distance $d = 0.34$ nm between the base pairs. A more compact form (A form) with about 11 bp per turn and $d = 0.29$ nm occurs at reduced water content. Certain DNA sequences with alternating g and c bases tend to form a left-handed helix (Z form), at least over a short distance, with 12 bp per turn and $d = 0.38$ nm. The different rotations give rise to a winding of the helix into coils (6). The *in vitro* structures of the DNAs listed in Table 3, however, are not known in detail (9), and the number of different structures is probably larger than the number of polytypes in layered compounds. The interactions between the bases of one strand are assumed to be very weak because of the absence of direct bonding (in helices without defects). Dipole–dipole interactions are considered besides the van der Waals interactions because of partial positive or negative charges in the neighborhood of the N–H– –O or N–H– –N hydrogen bonds. The present investigation shows that sequences of bases deviate from a random distribution ($\alpha_i = 0$) in such a way that attractive or repulsive interactions are preferred in some portions of the DNA.

2. CLOSE-PACKED LAYERED STRUCTURES

Many metals and alloys form close-packed structures. Metal atoms with weak bonding behave like tennis balls packed in a basket. The packing consists of coplanar hexagonal layers (Fig. 2). Successive layers are stacked, with centers of spheres falling directly above the centers of the triangular interstices of the two-dimensional layer below. Each layer has two sets of triangular interstices, as outlined

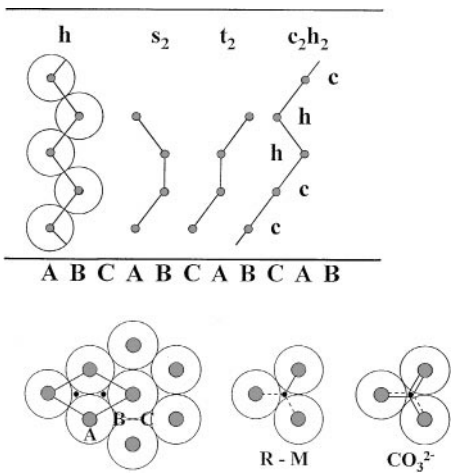

FIG. 2. Close-packed hexagonal layer of spheres in *A* position, unit cell ((0 0 1) projection, below) and *B* or *C* position of consecutive layers, e.g., *ABAB* (*hh*) or *BCAC* (c_2h_2), (*A*)*BB*(*A*) (s_2), and (*C*)*AA*(*B*) (t_2), as explained in the text ((1 1 0) projection, above). The single *R*–*M* bond is delocalized to three neighbors similar to CO_3^{2-} (below).

for the unit cell in Fig. 2. Each of the two sets forms the same lattice as the spheres themselves. The three different positions, which can be occupied by metal atoms in successive layers, are conventionally denoted as the *A*, *B*, and *C* positions. Most of the elements except Eu, U, and Np can be obtained in a close-packed modification (1, 4). Some metals like La and Ce or noble gases like He can crystallize in different structures depending on the temperature and on the isotope (He). The lanthanides La to Sm, Gd, and the actinides Am to Cf crystallize in the *ABAC* or $chch = (ch)_2$ stacking. Sm, Gd, and the heavier lanthanides Tb to Ho, Tm, Lu can crystallize also in the *ABABCBCAC* or $chhchhchh = (ch_2)_3$ stacking. The Jagodzinski–Wyckoff stacking symbols *h* and *c* describe a layer as *h* or *c* type according to whether the layers on each side of it are of the same type, such as for *B* in *ABA* or of different types as for *B* in *ABC*. The metals show a preference for simple sequences such as *hh*, *ccc*, $(ch)_2$, $(ch_2)_3$, or $(hc_2)_3$. Other sequences such as *hc* hc_4, *ch* ch_3, *hc* hc_2, *ch* ch_2, ch_6, ch_3, ch_2 ch_3, or h_2c_2 are found in alloys M_xN_y (1,4).

Atoms of close-packed structures with diameter $d = 1$ have $T_1 = 6$ nearest neighbors in the same layer and $T_1 = 6$ neighbors in the two adjacent layers (3 in each) at distance *d*. The number $T_2 = 6$ of second-nearest neighbors at distance $\sqrt{2}d$ is also identical for all close-packed structures. The atoms of the *hh* structure, however, have $T_3 = 2$ third-nearest neighbors at distance $\sqrt{8/3}d$ and 18 fourth-nearest neighbors at distance $\sqrt{3}d$, while atoms of *ccc* packing have 24 third-nearest neighbors at distance $\sqrt{3}d$. The metal atoms of the sequences *hc*, *chc*, *hch*, ch_3, and h_2c_3 have two different values of T_3 and higher coordination shells. The sequences *hc*, *chc*, and *hch* are observed in NiAs, NbS_2, and $CdCl_2$ structures, respectively. The more complex structures contain three or more atoms with different environments such as $hchc_4$, c_3h_4, or $ch_2c_2h_2$ (for MN_2R_4 composition), $chch_3$ (MN_2R_3), or hc_3, h_2c_6 or ch_2chc_2h (MNR_2). These stackings, however, are not observed for these compositions.

The structures of the $M_xN_yR_z$ compounds (1, 2) listed in Table 1a are more complex because of the different coordination of *M* or *N* metal atoms between layers of nonmetal *R* atoms. These compounds are usually described by the stacking of the *R* atoms and by a sequence of Greek or lowercase letters for *M* and *N* atoms at tetrahedral or octahedral positions (1, 2). In the present investigation all *M*, *N*, and *R* atoms are characterized by *h*, *c*, *s*, or *t* symbols to reduce the number of strings. The s_2 and t_2 groups with two atoms at the same positions (at different projection height) are defined similar to the *h* and *c* symbols: the *trans* (or t_2) configuration for different positions of the neighbors next to t_2 and the *cis* (or s_2) configuration at identical positions. In that case the different coordinations of *M* in a sequence of *R M R* atoms can be obtained by the symbol (*c* = octahedral, *h* = trigonal prismatic, *t* or *s* = tetrahedral), and structures with the same sequences of *R* and *M* atoms can be compared by the different coordinations of *M* atoms or —in other words—the different configurations of the chain (Table 2). The alternative description by sequences of *A*, *B*, and *C* positions is not specific and in many cases longer by a factor 2 or 3 than the present notation as was shown before for the $(ch)_2$ α-Nd and $(ch_2)_3$ α-Sm structures. The series of ZnS or SiC polytypes $(hc_2)_{2n+1}hc$ and $(hc_2)_{2n+1}hc_3$ at consideration of *R* atom packing (1) corresponds to $(s_2t_4)_{2n+1}\ s_2t_2$ and $(s_2t_4)_{2n+1}\ s_2t_6$ in the present notation ($h = s_2$, $c = t_2$). The combinations of ab_2 and *ab* structural units of the first series is on the right border of the structure map, while ab_2 and ab_3 combinations of the second series ($a = h$, $b = c$) are on S approaching ab_2 at large *n*.

The hexagonal or pseudohexagonal unit cells exhibit the approximate dimensions $a \approx 0.38$ nm, $c = 0.31pz$ nm for sulfides (2) (*z* is the number of *R* atoms per formula unit and $p = 1, 2$, or 3 is the number of chains to obtain identical *A* positions, such as $p = 2$ for $(ch)_2$ α-Nd structure). The distance between *R* atoms is usually increased by ~7% for *M* atoms in octahedral coordination and decreased by ~7% for *M* atoms in tetrahedral coordination. The ratio $\Delta z\ c/a \approx 0.82$ ($p = 1$, $z = 1$, Δz is the difference in height) corresponds to the close-packing of spheres (with a diameter $d = 0.38$ nm). The diameters of Se or Te atoms are ~6 or 12% larger in the isotypic compounds. These distances are much larger than the Se–Se or Te–Te distances 0.23 or 0.29 nm in the chains of hexagonal Se or Te. These chains are arranged spirally about the *c* axis with a nearest approach by 0.35 or 0.37 nm of Se or Te atoms in different chains. The density of Se or Te atoms in the helix structures is decreased by ~40% compared to the close-packed structures.

TABLE 2

Composition, Space Group (SG), Pearson Symbol (PS), Layer, and Stacking Sequences of $M_xN_yR_z$ Layered Compounds with the *h*, *c*, s_2, t_2 Notation (Fig. 2)

$M_xN_yR_z$	SG	PS	Layer sequence	Stacking sequence
RM, *RMR*, $(RM)_2R$, or $(RM)_3R$ chains (*I* form with $0.2 \le \Delta z \cdot c/a \le 0.8$)				
Zn S 3 C	216	cF8	*RM*	t_2
Zn S 2 H	186	hP4	*RM*	s_2
Zn S 4 H	186	hP8	*RM*	$s_2\ t_2$
Zn S 6 H	186	hP12	*RM*	$s_2\ t_4$
Zn S 12 R	160	hR24	*RM*	$s_4\ t_4$
Na Cl	225	cF8	*RM*	c_2
Ni As	194	hP4	*RM*	*hc*
δ' NNb	194	hP4	*RM*	*hc*
TiS HT	166	hR18	*RM*	$hchcc_2$
Mo S_2 2 H	194	hP6	*RMR*	h_3
Cd I_2 C 6	164	hP3	*RMR*	c_3
C Ta_2	164	hP3	*RMR*	c_3
Nb S_2 2 H HT	194	hP6	*RMR*	*chc*
Cd Cl_2 C 19	166	hR9	*RMR*	*hch*
Cd I_2 C 27	186	hP6	*RMR*	hc_2
N(ScTa)	186	hP6	*RMR*	hc_2
Mo S_2 3 R	160	hR9	*RMR*	ch_2
Ta Se_2 4 H (b)	194	hP12	*RMR*	c_3chc
Th I_2 β	194	hP12	*RMR*	h_3hch
Nb Se_2 4 s (d2)	187	hP12	*RMR*	$c_3h_3c_3chc$
Nb Se_2 4 H LT	187	hP12	*RMR*	$chch_2cchcch_2$
Ta Se_2 4 s (c)	186	hP12	*RMR*	$chc\ ch_2$
Nb Se_2 4 s (d1)	156	hP12	*RMR*	$ch_2c_2hchc\ c_3$
Ta S_2 6 s	160	hR18	*RMR*	$ch_2\ c_3$
Sn S_2 18 R	166	hR27	*RMR*	$c_3hc_2\ chc$
Bi_2 Te_3 I	166	hR15	*RMRMR*	c_5
Pt_2 Sn_3	194	hP10	*RMRMR*	*hchch*
In_2 S_3 γ	164	hP5	*RMRMR*	hc_3h
C_3 Al_4	166	hR21	*RMRMRMR*	hs_2cs_2h
Th_3 N_4	166	hR21	*RMRMRMR*	hc_5h
As_3 Sn_4	166	hR21	*RMRMRMR*	c_7
C_3 V_4	166	hR21	*RMRMRMR*	c_7
$RN(RM)_n\ RNR$ chains				
Sn Sb_2 Te_4	166	hR21	*RNRMRNR*	c_7
In In_2 Te_4 HP	166	hR21	*RNRMRNR*	c_7
Fe Fe_2 S_4	166	hR21	*RNRMRNR*	*hchchch*
Mg Al_2 Se_4	166	hR21	*RNRMRNR*	hs_2cs_2h
(Cr,Fe) $(Ga,Fe)_2Se_4$	166	hR21	*RNRMRNR*	ct_2ct_2c
Fe Ga_2 S_4 α	164	hP7	*RNRMRNR*	ht_2ct_2h
Pb_2 Bi_2 Se_5	164	hP9	$RN(RM)_2RNR$	c_9
Fe_2 Ga_2 S_5 2H	194	hP18	$RN(RM)_2RNR$	ht_2chct_2h
$Fe_2Ga_2S_5$ 3R	166	hR27	$RN(RM)_2RNR$	$ht_2c_3t_2h$
(Mn_2In_2) Se_5	166	hR27	$RN(RM)_2RNR$	$ct_2c_3t_2c$
$Mg_2Al_2Se_5$	164	hP9	$RN(RM)_2RNR$	$hs_2c_3s_2h$
$Ge_3As_2Te_6$	166	hR33	$RN(RM)_3RNR$	c_{11}
$Ge_4As_2Te_7$	166	hR39	$RN(RM)_4RNR$	c_{13}
$Ge_5As_2Te_8$	164	hP15	$RN(RM)_5RNR$	c_{15}
$RN(RM)_n\ RNR(MR)_m\ NR$ chains				
N C_3 Al_5	186	hP18	*RNRNRMRNR*	hs_2cs_4h
N_2 C_3 Al_6	166	hR33	*RNRMRNRMRNR*	hs_4cs_4h
N_3 C_3 Al_7	186	hP26	$RN(RM)_2RNRMRNR$	hs_6cs_4h
N_4 C_3 Al_8	166	hR45	$RN(RM)_2RNR(MR)_2NR$	hs_6cs_6h
Combinations of chains				
Zn In_2 S_4 IIb	186	hP14	*RM RNRNR*	hs_2ct_2h
Zn In_2 S_4 IIIa	160	hR21	*RM RNRNR*	hs_2cs_2h
Zn In_2 S_4 I	156	hP7	*RM RNRNR*	ht_2ct_2h
Zn In_2 S_4 VIa	160	hR42	*RM RNRNR*	$hs_2ct_2h\ ht_2cs_2h$
Zn In_2 S_4 IIa	164	hP14	*RNRNR MR RM RNRNR*	ht_2ct_2h

TABLE 2—*Continued*

$M_xN_yR_z$	SG	PS	Layer sequence	Stacking sequence
Combinations of chains				
Zn_2 In_2 S_5 IIa (1)	186	hP18	*RM RNRMRNR*	$ht_2ct_2s_2h$
$(Zn_2$ $In_2)$ S_5 IIa (2)	186	hP18	*RM RNRMRNR*	hs_2cs_4h
Zn_2 In_2 S_5 IIIa	160	hR27	*RM RNRMRNR*	$hs_2ct_2s_2h$
$(Zn_3$ $In_2)$ S_6 Ia	164	hP11	*RM RM RNRMRNR*	$hs_2t_2ct_2s_2h$
Zn_3 In_2 S_6 Ib	156	hP11	*RM RM RNRMRNR*	$hs_2t_2ct_2s_2h$
Ge Bi_4 Te_7	164	hP12	*RNRMRNR RNRNR*	c_{12}
chains containing vacancies at M' or N' positions				
$(Ga,In)_{3-x}$ InS_5	186	hP18	*RM′ RNRM′RM′R*	hs_2cs_4h
$(Zn,In)_{4-x}$ $(Se,S)_5$ IIb	164	hP18	*RN′RM′RNRN′R*	$hs_2t_2ct_2h$
			RN′RNRM′RN′R	$ht_2ct_2s_2h$
$(In,Ga)_{4.67}$ S_7	164	hP14	$RN'RMRN'R\ (RM')_2R$	$ht_2ct_2h\ hc_3h$
$(In,Ga)_{7.33}$ S_{11}	166	hR72	$RM'RMR\ N'R\ (RM')_2R$	$hs_2ct_2hhc_3h$
			RN′ RMRM′R	ht_2cs_2h
M–M bonding				
Zr Cl	166	hR12	RM_2R	c_4
Pt Te	12	mC8	RM_2R	c_4
Ga S 2H (β)	194	hP8	RM_2R	hs_2h
Ga Se 3R (γ)	160	hR12	RM_2R	cs_2h
Ga Se 4H (δ)	186	hP16	RM_2R	cs_2hhs_2h
Ga Se 2H (ε)	187	hP8	RM_2R	cs_2chs_2h
Pt_3 Te_4	166	hR21	$RM_2R\ RMR$	c_7
Pt_2 Te_3	166	hR30	$RMR\ RM_2R\ RMR$	c_{10}
Bi Se	164	hP12	$RMRMR\ RMRMR\ M_2$	c_{12}
Bi_4 Se_3	166	hR21	$RMRMR\ M_2$	c_7
Bi_2 Te	166	hR27	$RMRMR\ M_2\ M_2$	c_9
Bi II	166	hR6	M_2	c_2
Bi_2 Te_3 II	160	HR15	$M_2\ R_3$	c_5
–M–M– or *–M–N–* bonding (*N* form with $\Delta z \cdot c/a \approx 0.8$)				
LiRh	187	hP2	*MN*	h_2
CuPt	166	hR6	*MN*	c_2
α-Nd	194	hP4	$M\infty$	*ch*
Tb HP	194	hP6	$M\infty$	hc_2
α-Sm	166	hR9	$M\infty$	ch_2
MN_n, M_2N_n(B form with $\Delta zc/a \approx 0.2$, *bcc*)				
Cs Cl	221	cP2	*MN*	c_2
Ce Cd_2	164	hP3	MN_2	c_3
Bi F_3, Al Fe_3	225	cF16	MN_3	c_4
Na Tl	227	cF16	M_2N_2	c_4
Al Mn Cu_2	225	cF16	*MRNR*	c_4
Ag Sb Li_2	216	cF16	MNR_2	c_4
Al_2 Li_3	166	hR15	M_2N_3	c_5
Si Fe_2 HT	164	hP6	$MNMN_3$	c_6
$Ni_2\square Al_3$	164	hP5	*MRMRNR*	c_6
Si_2 Li_5	166	hR21	MN_2MN_3	c_7
Ga_4 Li_5	164	hP9	$M_2N_2M_2N_3$	c_9
Ge_2 Cu_2 Li_5	164	hP9	MNR_2NMR_3	c_9
Pb_3 Li_8	166	hR33	$MN_2MN_3MN_3$	c_{11}
Pb_2 Li_7	164	hP9	MN_3MN_4	c_{11}

Note. The distance between layers $\Delta z\ c/a$ varies in *N*, *I*, and *B* form.

The distance between layers $\Delta z \cdot c/a$ decreases to about 0.4 (0.2–0.8) for the occupation of interstitial sites (octahedral, trigonal prismatic or tetrahedral sites, I form, Table 2) and to $\sim$0.2 for the occupation of octahedral and tetrahedral sites or the body-centered lattice (B form, Table 2).

3. RULES FOR LINEAR VALENCE COMPOUNDS

In valence compounds $M_xN_yR_z$ the octets of the atoms can be completed if the sum of electrons is 8 or a multiple of 8 (1):

$$x\ e_M + y\ e_N + z\ e_R = 8z.$$

This equation satisfies the requirement of electroneutrality and can be applied for different distributions of valence electrons, e.g., in ZnS (Table 1):

(a) Formation of Zn^{2+} and S^{2-} ions (ionic chain).

(b) The eight electrons (four pairs) form four bonds to Zn atoms in tetrahedral coordination (1, 5).

(c) The electrons are shared between neutral S and Zn atoms with two covalent bonds (covalent chain).

Each atom in a close-packed layered compound has three neighbors in the next layer ($1 + 3$ for tetrahedral t_2 or s_2 configuration) (Fig. 2). One bond of –S–Zn– with S in t or s position is delocalized to three Zn atoms. Therefore these bonds are weaker than the remaining bond. The delocalization of the double bond of the S=Zn group enhances the bond strength to three neighbors compared to the fourth neighbor. The same bond strength for all Zn atoms is achieved at a consideration of both formulas –S–Zn–S– and S=Zn. The structural data show positive or negative deviations. Equal bonding to all neighboring atoms, as in (b), is also achieved for the three M atoms in octahedral or trigonal prismatic coordination if the chain is symmetric as in –S–Mg– or at consideration of resonance hybrids, e.g., of Cu I–Cu and Cu–I Cu or Mg S=Mg and Mg=S Mg. The covalent chain (c) allows close-packed layers of neutral atoms in the (0 0 1) plane and different sequences of M, N, and R atoms in (1 1 0), which can explain the ordering of these atoms in most structures.

Table 1a shows the different combinations for binary and ternary compounds. The R and M atoms usually alternate until an M atom is missing, which is indicated by $e_M = 0$ (5). Therefore chains of RMR occur in $0\ 2\ 7_2$, $(0\ 6\ 1_2)$, $0\ 4\ 6_2$, and $(0\ 4\ 2_2)$ compounds and chains with three or four R atoms in the remaining structures containing 0. The $(0\ 6\ 1_2)$ or $(0\ 4\ 2_2)$ structures are antistructures to $0\ 2\ 7_2$ or $0\ 4\ 6_2$ with $e_N + e_{N'}$ or $e_R + e_{R'} = 8n$ (1). In normal valence compounds M_xR_z without M–M or R–R bonding the number of bonds pointing to each M or R atom is proportional to the number of R or M atoms (z or x). The R=M–R–M=R chains of M_2R_3 compositions are obtained for divalent R and trivalent M (tetravalent R and hexavalent M must not be considered). The M_3R_4 composition yields the largest chain with three tetravalent M atoms. In ternary compounds the length of the chains is extended if 2 6 groups are introduced into the $0\ 3_2\ 6_3$ chain or 3 5 groups into the $0\ 4_3\ 5_4$ chain at the single bond position. Table 2 shows many experimental examples of $RN\ (RM)_n\ RNR$ and $RN\ (RM)_n\ RNR\ (MR)_m$ NR chains. The last example with two different possibilities for the introduction of –N=Al– groups was only observed for the antistructures with composition $N_xC_3Al_{4+x}$ ($x = 0$–4). The chains are symmetrical for $x = 2$ and 4.

Another theoretical approach is the combination of different chains such as S=Zn S=In–S–In=S in $ZnIn_2S_4$ modifications, or Te=Bi–Te–Ge–Te–Bi=Te Te=Bi–Te–Bi=Te in $GeBi_4Te_7$ with divalent Ge (1) (Table 2).

Some compounds with trivalent Ga and In contain ~30% vacancies at the M' or N' positions (Table 2). In that case ~1 of the 3 M or N atoms in the neighboring layers is missing and the average charge of M or N atoms is reduced to ~2. This can explain the occurrence of vacancies in some cases, e.g., the first Ga′ and the Ga″ position in $(Ga,In)_{3-x}$ InS_5 with S=Ga′S=In–S–Ga″–S–Ga′=S (Ga′=$Ga_{0.50}In_{0.17}$, Ga″=$Ga_{0.67}$). A second formulation S=Ga′ S=In–S–Ga″=*S* Ga′=*S* with equivalent Ga′ positions requires no vacancies in the Ga″ position. The $(In,Ga)_{4.67}S_7$ and $(In,Ga)_{7.33}S_{11}$ compounds contain R–M–R–M–R units of In_2S_3 structure with hc_3h stacking, but ~0.2 additional M atom next to the R atoms at the end of the chain (1) (the Pearson symbols are hP14 and hR72 instead of hP12 and hR57).

The last group of Table 2 considers general valence compounds. These are valence compounds either where the cations do not transfer all their valence electrons—they are used for bonds between them or for nonbonding orbitals—or where the anions due to bonds between themselves do not need as many electrons to complete their octet shell (1). For a compound M_xR_z with

$$x(e_M - e_{MM}) + z\ (e_R + e_{RR}) = 8z$$

the average number of remaining valence electrons e_{MM} and e_{RR} at M or R atoms can be used to form M–M or R–R bonds, respectively, as for Cl–Zr≡Zr–Cl, Te=Pt=Pt=Te, or S=Ga–Ga=S in Zr Cl, Pt Te, Ga S or Ga Se structures. Bi_2Te_3II (–Bi=Bi–Te–Te–Te–) and the hexagonal structures of Se and Te mentioned before are examples for –Se–Se– or –Te–Te– bonding. The Pt_3Te_4 and Pt_2Te_3 structures can be considered combinations of Te=Pt=Pt=Te and Te=Pt=Te chains. Pt and Zr are assumed to be tetravalent, though they do not belong to the elements considered for valence compounds (1). These atoms with lone electron pairs or low electronegativity (Mg, Zr), given in parentheses (Table 1b), are ordered for increasing size (10). The Zn, Al, Ga, and N atoms underlined in Table 1b prefer tetrahedral sites. Divalent Mn, trivalent Fe or In, and tetravalent C (dashed underlined) are observed in tetrahedral or octahedral coordination. Besides this, some ordered alloys such as Li Rh, Cu Pt, and Pt_2Sn_3 and some metallic carbides or nitrides such as Ta_2C or Th_3N_4 (which are not valence compounds) are also listed in Table 2 because of related structures. The distance between layers $\Delta z \cdot c/a \approx 0.8$ is normal (N form) in the close-packed metals, decreased to

0.2–0.8 at occupation of interstitial sites (I form), or to ≈ 0.2 in ordered body-centered alloys (B form). B_4Rh_5 with the sequence *RMRMRMRMR* (1), not allowed in the present scheme, is omitted in Table 2. The distances $\Delta z\ c/a$ (at $p = 1$, $z = 1$) between neighboring layers usually correspond with the bond strength. The distances between Th and N layers, e.g., of Th_3N_4 (N≡Th–N=Th=N–Th≡N sequence) are 0.51 (Th–N), 0.31 (Th=N), 0.12 (Th≡N) and 0.49 (N to N of next chain).

4. STRUCTURE MAPS OF LAYERED COMPOUNDS

The *M*, *N*, and *R* atoms of the different chains listed in Table 2 and the sequences of a, t, g, and c bases (11) were analyzed for the different self-coordination numbers and the different α_i values plotted in structure maps as was outlined in Section 1.

The layer sequences *MR* (*ab*), *RMRMRMR* ($abab_2$), *RMRMR* ($abab_2$), *RMR* (ab_2), *RMR RMMR RMR* ($ab_2\ a_2b_2\ ab_2$), *RMMR RMR* ($a_2b_2\ ab_2$), and *RMMR* (a_2b_2) (with $M{=}a$) are at the right border of the map. The interactions between *M* atoms are repulsive in *MR* and attractive in *RMMR* (a_2b_2) and are assumed to vary with increasing α_2. The positions of the *M* and *R* atoms of these binary compounds are at the limit of the straight lines C (cluster), S (single), D (double), T (triple), and M (multi) units. All ab_n chains with $n \geq 2$ and combinations of different chains such as $ab_2\ ab_3$ are on the line S. All $abab_n$ containing two alternating *a* are on line D. The combinations of three alternating *a* in $abababab_n$ are on T and alternating *ab* units at M (multi units). At the left border of the map (line C) the size of the a_n and b_n clusters increases from a_2b_2 to complete segregation $a_\infty b_\infty$ at increased α_2.

The *M* atoms of ternary chains such as $RN(RM)_nRNR$ or *RM RN RNR* are on the lines S, D, or T depending on the number of *RM* units. A layered quasi-crystal with a Fibonacci sequence of ab_2 and *ab* units ($ab_2/ab \approx 1.62$) corresponds to QC ($\alpha_2 = 0$). The *M* and *N* atoms of $(Zn_2In_2)S_5$ IIa (2), $(Ga, In)_{3-x}InS_5$, and other compounds listed in Table 2 (with *M* and *N* in brackets) are disordered. In that case the positions of *M* and *N* approach $\alpha_i = 0$ as for disordered alloys (4).

5. NUCLEIC ACIDS

The primary structure of nucleic acids is given by the sequences of the bases adenine (a), thymine (t), cytosine (c), and guanine (g). Table 3 shows the number of a, t, c, and g units to calculate, e.g., the ratio $r = (t + c + g)/a$ for the base a. The ratios $r = 1.5$–7.5 of the nucleic acids listed in Table 3 are similar to those for the layered compounds ($r = 1$–11). The $\alpha_1\ \alpha_2$ values which were obtained from the averaged self-coordination numbers and *r* values (Sect. 1) are on S or S′ (negative S) or in the fields C1-C4 or D1-D4 with α_i values similar to those in disordered alloys (4). An alternation of different units such that identical units are separated by at least two other units yields $-\alpha_{1,2} = 1/r$ on line S with the minimum value $r = 2$ in ab_2 (Fig. 1). At $r < 2$ (as, e.g., *a* in secA 7 or 10) the bases can cluster or alternate with other bases (ab). The maximum and minimum α_i values

TABLE 3
Characterization of Nucleic Acids by the Number of Adenine (a), Thymine or Uracil (t), Cytosine (c), and Guanine (g) Units (First Number), the Clusters per Mill (Second Number) and the Location on the Structure Map

Nucleic acid	a	t	c	g
secA 7 *S.carnosus*	1308/154/C3	990/75/C3	564/7/C1	727/21/C1
secA 10 *O.lutens* chloroplast	1212/207/C3,4	980/169/C3	359/25/C3,4	508/24/C3
enolase 3 *S.cerevisiae*	523/98/C3	634/120/C3	443/29/C1,2	413/24/C1
enolase 59 *A.glutinosa* mRNA	473/99/C3	438/71/C4	329/33/C1	415/39/C1,2
enolase 249 *Schisosoma mansoni* mRNA	12731/125/C4	12776/127/C4	8117/35/C2	8040/37/C2
enolase 2 *N.fontalis*	691/188/C4	613/90/C2	303/36/C4	361/19/C2
enolase 100 *Streptococcus thermophilus*	2323/195/C4	1933/153/C4	1021/36/C3	1222/46/C3,4
enolase 1 *M.musculus* mRNA	363/72/C3	285/28/D2	378/56/C2	407/91/C3
enolase 8 *H.sapiens* mRNA	340/50/C3	271/22/D1	367/68/C3	420/114/C3
enolase 15 Rat norwegiens mRNA	396/61/C3	378/40/D2	490/84/C2	461/61/C2
enolase 13 Peking duck tan.cryst.	439/71/C3,4	415/58/D1,2	417/96/C3	444/43/C1
enolase 25 *Xenopus laevis* mRNA	461/93/C4	448/54/D1,2	370/62/C2	441/50/C1,2
secA 21 *Synechococcus* sp.	905/70/C3	777/51/C2	989/39/C1	1121/40/D1
secA 6 *P.sativum* mRNA	1120/124/C3,4	996/94/C4	688/28/C1,2	804/50/D3,4
secA 11 *S.oleracea* chloroplast	1084/101/C3,4	1043/82/C4	689/45/S′	923/53/D1,2
secA 15 *Mycobacterium bovis*	782/35/C2	667/21/C1,2	1286/38/S	1310/45/D1
secA 8 *Escherichia coli*	1016/121/C3	853/74/C3	938/29/D1	1004/27/S
secA 4 *Streptomyces lividans*	709/4/C1	509/10/D1	1351/53/C1	1437/76/D1

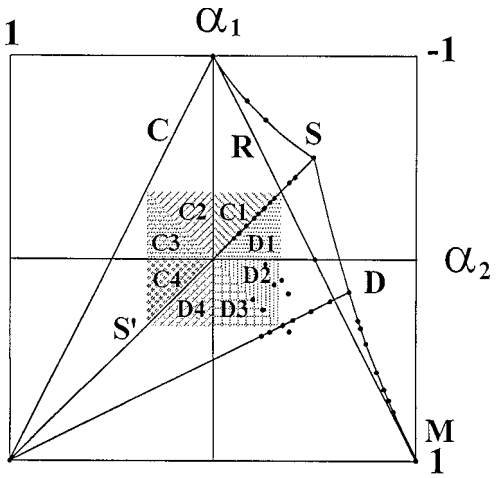

FIG. 3. α_1 α_2 structure map with experimental points of $M_xN_yR_z$ compounds (see text) and different fields C1–C4 and D1–D4 for a, t, g, and c bases of the nucleic acids listed in Table 3.

of $\sim \pm 0.2$ are smaller than the $1/r$ values because of the coexistence of clusters (C), single units (S), and double units (D). All sequences of the first group are in the field C1-C4 because of the increased number of clusters (Fig. 3). The structure map suggests small clusters such as a_2b_2 in C1 and larger clusters in C2, 3, or 4. The α_1, α_2 values of t units in the second group of eukaryotic (nonbacterial) enolases are in the field D1 or D2 with an increased number of double units (D). In the third group of secAs all α_1, α_2 values of g and some values of c contain an increased number of double units (D). The g value of secA 8 is on S. In secA 4 the t and g units have an increased number of double units (D). The four points at negative S (c in secA 11), C4 or D4 (t and g in secA 6) are with extended clusters besides alternating *ab* units. In the first group the percentage of clusters is increased to 10–21% (a) or 7–17% (t) or decreased to 1–5% for c and g compared to the average of 4% for t and 7% for a, c and g.

The mRNA carries the information for the sequence of amino acids in a protein in the form of the genetic code in which each occurrence of one of the 64 groups of three nucleotides (triplets or codons) conveys the information for a specific amino acid (or in some cases a stop signal to indicate the end of the protein) (6). Fifteen amino acids are obtained by symmetrical triplets aaa (Lys), ata (Ile), aca (Thr), aga (Arg), ttt (Phe), tct (Ser), tgt (Cys), tat (Tyr), ccc (Pro), cgc (Arg), cac (His), ctc (Leu), ggg (Gly), gag (Glu), gtg (Val), and gcg (Ala). Arginine (Arg) is obtained by aga or cgc.

The remaining five amino acids, Asn, Asp, Glu, Met, and Trp, can be obtained only by nonsymmetrical triplets. This is only possible because of a given direction from 5′ to 3′ on each strand (6) and of stop signals (taa, tag, or tga).

The increased fraction of a and t clusters in the first group would be obtained, e.g., at an increased number of aaa (Lys) and ttt (Phe) clusters, while the increased percentage of double units for t in the second group could be possible at an increased number of tct (Ser), tgt (Cys), and tat (Tyr).

Two types of right-handed helices (A or B form) with 11 or 10 bp per turn and a left-handed helix (Z form) are known (7). Within the cell the DNA helix is wound up into coils (6). For example, bacterial DNA is normally negatively supercoiled to the extent of about one negative turn per 200 bp. The density is increased by the supercoiling. The *Escherichia coli* chromosome, in its expanded state, for example, would be several hundred times longer than the bacterial cell itself (6).

There are six parameters for the exact analysis of the translation and rotation from one base pair to the next, which are correlated and depend to a certain extent upon the choice of overall helix axis (8). An approximate description with A, B or Z indices for each base pair in the A, B or Z helix and X values for deviations from the ideal helix by about 10% would show whether there is a correlation between the sequence of bases with, e.g., clustering in certain areas and the approximated configuration of the helix. The sections of the helix containing single bases ($a + g \neq c + t$) could be indicated by A′, B′, Z′, and X′, respectively.

6. CONCLUSIONS

The layered structures or the nucleic acids can be described by sequences of three or four units. The layered structures with the composition $M_xN_yR_z$ are approximately close-packed R atoms with M and N atoms at interstitial sites. The 74% density of the sphere packing is increased at complete occupation of all the tetrahedral and octahedral interstices by 1.7 and 5.3%, respectively (3, 4). The compounds are obtained by divalent $M = Zn$, Mg, Fe, Mn, trivalent $N = Al$, Ga, In in sulfides or selenides, or the antistructures with composition $N_xC_3Al_{4+x}$. These two cases can be considered delocalized covalent bonding within the chains S=Ga–(S–Fe–S–)$_x$Ga=S for $Fe_xGa_2S_{4+x'}$ such as GaS,$FeGa_2S_4$,$Fe_2Ga_2S_5$ or in chains Al≡C(–Al=N)$_{x-1}$–Al=C=Al–(N=Al–)$_{x-1}$C≡Al in $N_xC_3Al_{4+x}$ ($x = 0, 2, 4$). Other structures with nonsymmetrical chains at insertion of additional units or exchange of M and N atoms are also possible (Table 1). The trivalent Fe, Al, Ga, and In, the divalent Mn or Zn, and the C and N atoms of $N_xC_3Al_{4+x}$ can be on tetrahedral sites. The tetrahedral sites of N or M atoms are characterized in the present investigation by s_2 or t_2 in the stacking sequence, if the neighboring atoms are on the same or on opposite sides (Fig. 1). In that case the idealized structures can be described by the layer sequences of M, N, R and the stacking sequences h, c, s_2, and t_2 (Table 2).

In many cases of polytypes such as for ZnS or $NbSe_2$ the stacking sequences are much larger than the layer

sequences. In some extended sequences of layers with an occupation of octahedral sites as for $Ge_xAs_2Te_{3+x}$ ($x = 3$–5) the stacking of all atoms is *c*. Most stacking and layer sequences are symmetric. In NC_3Al_5 and $N_3C_3Al_7$ both sequences are asymmetric. Few cases like Cd I_2 C27 or MoS_2 3R with *RMR* layer sequence exhibit asymmetric stacking sequences (hc_2 and ch_2). Some polytypes of this group such as TaS_2 6s with ch_2c_3 can be considered a combination of ch_2 (MoS_2 3R) and c_3 (Cd I_2 C6). The five stackings *hc* (NiAs), *chc* (NbS_2), *hch* (Cd Cl_2), ch_3 (*MR*), and h_2c_3 (MR_4) exhibit identical T_i values of all *M* and *R* atoms in the three-dimensional packing and can be compared with the Cu or Mg structure with a single set of T_i values (3, 4). The NiAs, NbS_2, and $CdCl_2$ structures are related to α-Nd, Tb HP, and α-Sm structures (1). The geometry of the $M_xN_yR_z$ compounds can be described approximately by the packing of *R* atoms with slightly increased or decreased distances at occupation of octahedral or tetrahedral interstices by the metal atoms *M*, *N*.

The present investigation shows that most of the layered $M_xN_yR_z$ structures can be described by chains of atoms with one to four directional bonds such as S=Ga–S–Fe–S–Ga=S, S=Sn=S, or Cl–Cd–Cl. The three-dimensional structures (Fig. 1) and the high symmetry of the chains (with threefold or sixfold axes) show that the bonds are delocalized within the three neighbors of the succeeding layers similarly to, e.g., the C=C double bond in benzene or the four bonds in CO_3^{2-} groups. The bond direction is fixed only for one neighbor at tetrahedral site occupation. The occurrence of so many polytypes with different stacking suggests a weak directional covalent bond coexisting with the ionic state (Table 1a). The formulation S≡Ga–S≡Fe≡S–Al≡S, S≡Sn≡S or Cl≡Cd≡Cl with six bonds for atoms in octahedral or trigonal prismatic coordination and four bonds for tetrahedral coordination (2) suggests a stronger bonding for atoms on octahedral sites. The formulas of the same examples given above (Tables 1a, 2) show, e.g., two bonds for divalent Fe and three bonds for trivalent Ga and can explain the reduced distance between the S atoms next to Ga and the increased distance for the S atoms next to Fe. An Ising-type analysis of the interactions between the different atoms shows a variation from repulsive interactions *RM* (*ab*) to attractive in *RMMR* (a_2b_2) sequences. The repulsive interactions decrease for *RMRMRMR* ($ababab_2$), *RMRMR* ($abab_2$), *RMR* (ab_2) and switch to increasing attractive interactions for *RMR RMMR RMR* ($ab_2\ a_2b_2\ ab_2$), *RMR RMMR* ($ab_2\ a_2b_2$), *RMMR* (a_2b_2) (with $M = a$) at the right border of the structure map (Fig. 1). The ionicity with a repulsion of charged atoms increases at increased electronegativity difference and:

(1) The *c* stacking, e.g., of MgO, CaS, or BaS in NaCl structure is favored.

(2) The t_2 compared to s_2 stacking is favored.

(3) The ionic chain (Table 1a) becomes unstable with respect to the three-dimensional distribution of *M* and *R* atoms, e.g., in $TiCl_3$ or NbO structures (4).

(4) The distance *a* between identical atoms of each layer is increased for $\Delta z\ c/a \approx 0.2$. The compounds are body-centered alloys, which are ordered in (1 1 1) layers (*B* form, Table 2). Compounds like BiF_3 are considered cubic close-packed Bi with F atoms at octahedral and tetrahedral interstices.

The nucleic acids are characterized by chains of 10^4–10^{12} bp containing nonperiodic sequences with 10–5000 bases a,t,c and g in two complementary strands of the double helix. The a and t bases or c and g bases are linked by two or three hydrogen bonds, but not with the neighboring bases at the same strand (unless there are defects). The bases of one strand are connected indirectly by the chain of alternating 2′-deoxyribose and phosphate groups. The long strands are characterized by the concentration of the bases and the averaged self-coordination numbers to calculate the short-range order parameters α_1 for nearest and α_2 for next-nearest neighbors within the chain. The deviation of the $\alpha_{1,2}$ values from a random distribution ($\alpha_i = 0$) can be analyzed for a short-range order structure similar to that in disordered alloys or disordered $M_xN_yR_z$ compounds. The map of $\alpha_1\ \alpha_2$ values can be analyzed for a preference of clusters a_2b_2 or a_3b_3, etc., single units ab_2, ab_3, etc., and double units $abab_2$, $abab_3$, etc., where *a* is the analyzed base and *b* the remaining three bases. The 18 DNA or RNA of the present investigation can be classified into four groups with different fields of the structure map. All bases of the first group are on C1-C4 fields of the structure map with an increased clustering of all bases. The t bases of the second group, the g bases of the third group and t or g of the last example (secA 4) contain an increased amount of double units $txtx_n$ or $gxgx_n$, where t or g bases are separated by $n \geq 2$ other bases x.

The different α_i values can be correlated with the percentage of aa clusters and to some extent also with the triplets of protein coding.

A sequence of A, B, Z, and X values for the stacking of the four bases similar to the *c*, *h*, *s*, and *t* values of the layered compounds (Table 2) would show the correlations between the sequences of bases and the approximated configuration of the helix (A, B, and Z forms).

REFERENCES

1. E. Parthé, L. Gelato, B. Chabot, M. Penzo, K. Cenzual, and R. Gladyshevskii, *in* "Gmelin Handbook of Inorganic and Organometallic Chemistry, TYPIX," Vol. 1. Springer-Verlag, Berlin, 1993.
2. F. Hulliger, "Physics and Chemistry of Materials with Layered Structures" (F. Lévy, Ed.) Vol. 5. Reidel, Dordrecht, 1976.
3. L. Pauling, "The Nature of the Chemical Bond," Cornell Univ. Press, Ithaca, NY, 1945,1960.

4. J. Hauck and K. Mika, *in* "Intermetallic Compounds," Vol. 1, "Principles" (J. H. Westbrook and R. L. Fleischer, Eds.), p. 277. Wiley, London, 1994.
5. J. Hauck and K. Mika, *J. Solid State Chem.* **138,** 334 (1998).
6. J. W. Dale, "Molecular Genetics of Bacteria," Wiley, Chichester, 1996.
7. J. P. Glusker, L. Mitchell, and M. Rossi, "Crystal Structure Analysis for Chemists and Biologists," VCH, Weinheim, 1994.
8. R. E. Diekerson *et al*., *J. Mol. Biol.* **205,** 787 (1989).
9. S. R. Holbrook and S.-H. Kim, *Biopolymers* **44,** 3 (1997).
10. K. Schubert, "Kristallstrukturen zweikomponentiger Phasen" [Crystal Structures of Binary Phases], Springer-Verlag, Berlin, 1964.
11. GenBank of the National Center for Biotechnology Information.